Alonso, C. and Romero, A. (2016). Wage Expectations for Higher Education Students in Spain. LABOUR, v. 3, n. 1 pp. 1-17. Available in: https://doi.org/710.1111/labr.12072

Wage Expectations for Higher Education Students in Spain

César Alonso-Borrego — Antonio Romero-Medina

Abstract. We assess students' ability to forecast future earnings by using data on expected wages self-reported by college students with different graduation horizons. We find a significant gender gap, by which wage expectations are systematically lower for women than for men. However, women do not fully account for the gender gap in their future earnings. We also find that student performance, degree type, and graduation horizon play a relevant role in wage forecasts. In any case, students' expectations do not conform market wages but become more realistic as they approach graduation.

1. Introduction

The accuracy of income expectations by students and its relation to education decisions is at the heart of the 'human capital' model. Wage expectations influence individual choices on education, investment, and labor supply. The purpose of this study is to compare earnings expectations of college students with the earnings reported by recent college graduates. This exercise will allow us to determine whether specific individual characteristics can be identified as potential sources of error in expectation formation process.

Furthermore, the availability of subjective expectations allows to learn about the process of expectation formation, and to improve our understanding of individual behavior (Dominitz and Manski, 1996). Our main findings are that both gender and past academic performance play a key role in the ability to form wage expectations. Furthermore, it is rather reassuring that student predictions become more realistic as students approach graduation. Provided that earnings expectations accurately estimate future earnings, we can argue (see Chevalier, 2007) that they pose the advantage of being free from some types of endogeneity. In particular, individuals already working and reporting low earnings might claim to be less financially motivated than similar individuals with higher earnings.

Despite its importance, the number of studies that assess the accuracy of income expectations is small, and the evidence is mixed. This paper contributes to the debate by using individual data on college students who report the wage they expect to earn after completing college. Comparing data on wage expectations with average actual wages for young college

We are in debt with Juanjo Dolado and Ricardo Mora for valuable comments. We acknowledge research funding from the Spanish Ministry of Finance and Competitiveness, Grant Nos. ECO2012-31358 (first author) and ECO2014-57442-P (second author).

Department of Economics, Universidad Carlos III de Madrid, 28903 Getafe, Spain. E-mail: aromero@eco.uc3m.es.

graduates, we can assess to what extent self-reported measures of expected wages are realistic, and how they depend on their graduation horizon. As far as the optimal education level and the degree choice are linked to the market returns to education (Betts, 1996), the results may shed light on how education choices are ruled.

We have used in our analysis demographic individual data, character trail information, as well as life ambition data. The main data for this study were gathered in Madrid from 2001 to 2005 among students in different public universities of the region. We also considered, as a benchmark, actual wages by gender for young college graduates employees, between 25 and 29 years old, working in the Madrid region. The main data offer two advantages to our purpose. First, we have students following different degrees, where the same degree can be followed in different colleges, which allows for broader comparisons. Second, all the colleges belong to a single region, which is a largely populated and an economically dynamic region that attracts workers from all over the country. Thus, we can assume that they have the regional market as their implicit target.

We estimate the model for expected wages, considering two different subsamples according to the time horizon for degree completion. Namely, we consider first-year college students and penultimate-year or junior college students. Our data set contains information about the degree and academic year for each student, as well as gender, pre-university and college academic performance, and socioeconomic background. We also include individual information by each student before entering university and additional reasons behind their degree choice.

We find that expected wages for fresher students are affected by gender. Furthermore, poor academic performance (before college and in college), the degree type, degree preferences, and household characteristics affect wage expectations. Expected wages predicted by the empirical model suggest a high level of overprediction with respect to actual observed wages. In the case of junior students, the poor academic performance plays a relevant role in determining expected wages. However, neither pre-university performance nor household environment influences their wage forecast. In general, most junior students exhibit a positive gap between expected and actual wages, but such gap tends to decrease as the students' horizon to graduation approaches.

Among the previous contributions, we should mention Dominitz and Manski (1996), Das and van Soest (1999), and more recently Brunello *et al.* (2004), Webbink and Hartog (2004), Delaney *et al.* (2010), and Jerrim (2011). The results show how differences in expectations may arise because of the particular degrees or particular colleges in which the sample is conditioned, differences in the sample sizes, etc. In particular, Brunello *et al.* (2004), using survey data from business and economics university students across 10 European countries, find that older individuals as well as senior students report lower earnings expectations. Webbink and Hartog (2004), using longitudinal data on higher education students, find that systematic under- or overestimation seldom manifests. Jerrim (2011) finds that full-time college students in the UK usually overestimate their starting salaries. Along the same path, Carvajal *et al.* (2000) find that in general students' expectations are in line with the salaries of recent graduates. Finally, Alonso-Borrego *et al.* (2007) exploit our same data to assess the students' expected and shadow wages. Expected wages predicted by their model are very similar across gender, whereas shadow wages are greater for female. Their statistical model consisted of an ordered response model, which, unlike our paper, ignores the quantitative information on wage thresholds. The fact that the thresholds determining the different ordinal categories are known allows us to implement a pointwise censored model, which avoids the need to formulate strong assumptions about the left and the right tails of the distribution to project

individual expected wages. We find, indeed, that the predictions of individual expected wages are very sensitive to these additional assumptions.

The remainder of the paper is organized as follows. In Section 2 we outline the data sets, the variables, and alternative model specifications. Sections 3 and 4 present the econometric framework and our estimation results. Section 5 provides some concluding remarks.

2. Data

2.1 The survey

The primary source of data is a survey financed by the Madrid regional authority and carried out in the academic years 2000–01, 2003–04 and 2004–05. The survey explored attitudes and opinions with regard to the higher education system of young students registered in public universities in the Madrid region.[1] This data set is particularly adequate because it allows to test the role of several variables that are potentially relevant for wage forecasts, i.e. gender, performance, and the time horizon to graduation.

Our data set provides information regarding wages expected after graduation. Each student is asked his/her expected monthly wage after concluding his/her studies: 'What is the monthly wage that you are expecting after graduating?' This question is similar to the ones formulated in other studies (see Jerrim, 2011, for a brief discussion about this matter). Taking into account the specificities of the job market in Spain, we expect the students to respond the gross monthly income in real terms that they expect to earn from their first full-time job after they graduated. Among the 1,659 students surveyed from all public universities in Madrid, we had 288 who did not answered or answered 'Don't know'. The answers provided by 1,371 students were surveyed into five discrete categories. In Table 1, we show the marginal relative frequencies of expected wages for each wage category in our sample. Expected wages exhibit a remarkable unimodal profile, whereby 53 per cent of students chose the third category (between 1,803 and 3,606 euro per month). We also find that a sizeable proportion of respondents reported expected wages in the highest, unbounded category (more than 5,409 euro).

The data set also contains information on gender, academic and personal status, and socioeconomic background of each student. It also provides specific information about the characteristics of the degree chosen by the student, as well as on the choice motivation and the alternative degrees considered. Finally, we have some information related to individual performance. In Table 2, we provide the statistics of the main variables.

Table 1. Monthly expected wages of Madrid college students

Relative frequency (%)	**All**	**First year**	**Junior**
Between 450 and 901 euro	4.9	2.6	6.3
Between 901 and 1,803 euro	17.9	12.7	22.2
Between 1,803 and 3,606 euro	52.7	47.5	55.6
Between 3,606 and 5,409 euro	13.8	18.7	9.9
More than 5,409 euro	10.8	18.4	6.0
Number of non-missing observations	1,371	385	284
Number of missing observations	288	111	58

Source: Young people facing college education, 2001, 2004, and 2005.

Table 2. Main variables and descriptive statistics

	All		Female		Male	
Variable	Mean	SD	Mean	SD	Mean	SD
Female	0.57	0.50				
Family						
High family income	0.19	0.39	0.14	0.35	0.26	0.44
University father	0.41	0.49	0.40	0.49	0.42	0.49
Pre-university						
Public secondary	0.58	0.49	0.59	0.49	0.57	0.49
Science secondary	0.52	0.50	0.50	0.50	0.54	0.50
Access grade	67.78	9.32	67.69	9.50	67.90	9.07
Examination passed at first attempt	0.84	0.37	0.87	0.34	0.81	0.40
Choice reasons						
Economic independence	0.66	0.47	0.63	0.48	0.70	0.46
Vocation	0.76	0.43	0.74	0.44	0.78	0.42
Parental influence	0.40	0.49	0.38	0.48	0.42	0.49
Difficulty	0.39	0.49	0.39	0.49	0.40	0.49
Choice set						
Same degree	0.15	0.35	0.15	0.36	0.14	0.35
University degree chosen						
First choice	0.61	0.49	0.66	0.47	0.54	0.50
Long degree	0.80	0.40	0.79	0.41	0.81	0.39
Science and engineering (S&E)	0.35	0.48	0.28	0.45	0.44	0.50
Health	0.13	0.34	0.18	0.39	0.07	0.25
Educational (Educ.)	0.07	0.25	0.09	0.28	0.04	0.24
Social Sciences (Soc.Sci.)	0.27	0.44	0.29	0.45	0.25	0.43
Humanities (Hum.)	0.18	0.38	0.16	0.37	0.20	0.40
College performance						
Grant	0.17	0.37	0.18	0.39	0.15	0.35
Repeater	0.30	0.46	0.27	0.45	0.35	0.48
Satisfied	0.21	0.41	0.25	0.43	0.16	0.37
Working	0.18	0.39	0.18	0.39	0.18	0.39
Survey year						
2004	0.31	0.46	0.25	0.43	0.40	0.49
2005	0.56	0.50	0.61	0.49	0.50	0.50

Source: Young people facing college education, 2001 2004, and 2005.
Note: All the variables are binary except for access grade, which ranges between 50 and 100.

We explore the role of the degree choice, the choice motivation, and the alternative degrees considered. We have broken down the degrees by duration and type. We consider whether the first three choices featured the same degree in several colleges; about 15 per cent of the sample students considered the same degree offered in different colleges. The degree attended by the sample students was the first choice for 61 per cent of them. Long degrees clearly dominate, amounting to 80 per cent of the sample. Regarding choice motivation, vocation and economic independence feature among the main reasons offered by 70 per cent of students. Other reasons such as family influence and the difficulty of the degree were alleged by 40 per cent of students.

Two features are particularly relevant for our analysis: performance and gender. Student performance might affect the ability to forecast income after graduation. Less than 20 per cent were awarded a grant. About 30 per cent of the students have failed and repeated at least one

academic year. Finally, only 20 per cent of students reported to be satisfied with their studies, and less than 20 per cent are simultaneously working (full-time or part-time work) and studying.

We find different patterns in the variables under study by gender. Regarding family characteristics, the percentage of students belonging to high-income households is clearly lower for females than for males. Regarding degree types, science and engineering (S&E hereinafter) are dominated by men, whereas health and education are dominated by women. Regarding individual characteristics and performance, men and women differ in terms of academic performance. In particular, a higher percentage of women passed the access examination at their first attempt, and follow college degrees that were their first choices. Women also seem to perform better in college, with a higher proportion of grants awarded[2] and a lower proportion of individuals who are forced to repeat one academic year. Finally, women report to be satisfied with college to a higher extent than men.

2.2 Complementary data

In order to assess the quality of wage forecast by college students, we make use of actual wages of comparable individuals as a benchmark. For this purpose, we consider wages by gender for young college graduates employees, between 25 and 29 years old, working in the Madrid region, from the 2002 Survey of Wage Structure, carried out by the National Institute of Statistics.

In Table 3, we present the marginal distribution of actual wages for young college graduate employees, between 25 and 29 years old. We can see that the wage distribution in Madrid is shifted to the right with respect to the distribution at the national level. In line with expected wages by college students, the distribution for actual wages in Madrid exhibits a unimodal profile. However, there are differences with the distributions of expected wages. The distribution of wages expected by junior students is slightly shifted to the right with respect to the distribution of actual wages for young graduates in Madrid, yet they do not differ very much.

In Table 4, we present the average monthly wage for young college graduates employees, split by gender and by the occupation related with each degree type and length. Because this information is widely publicized and easily accessible, it is reasonable to assume that it is part of the information set that university students used when computing their expected wages. In this breakdown, we must note there are not short degrees in humanities.

The average wages in this complementary data set are representative of the population of employees.[3] Therefore, such information is potentially affected by two sources of selection bias. The first one is related to the decision on labor participation, which differs for women

Table 3. Monthly earnings for college graduates employees aged 25–29 years

Relative frequency (%)	National	Madrid
Between 450 and 901 euro	15.1	12.6
Between 901 and 1,803 euro	44.6	39.3
Between 1,803 and 3,606 euro	37.7	43.4
Between 3,606 and 5,409 euro	2.1	4.4
More than 5,409 euro	0.5	0.3

Source: Own calculations from 'National Survey of Wage Structure', 2002.

Table 4. Monthly average earnings (in euro) for college graduates employees aged 25–29 years, by degree type and gender

	S&E	Health	Soc.Sci.	Hum.	Educ.	Total
National average, long degree						
Male	2,113	1,732	1,702	1,482	1,533	1,859
	(896)	(731)	(930)	(731)	(626)	(882)
Female	1,912	1,542	1,566	1,538	1,483	1,628
	(985)	(565)	(829)	(680)	(570)	(793)
Total	2,043	1,616	1,624	1,516	1,505	1,742
	(931)	(641)	(875)	(699)	(593)	(846)
National, short degree						
Male	1,850	1,704	1,242		1,481	1,757
	(954)	(679)	(568)		(533)	(895)
Female	1,704	1,655	1,216		1,385	1,522
	(696)	(533)	(782)		(487)	(627)
Total	1,820	1,663	1,225		1,414	1,644
	(907)	(558)	(716)		(502)	(787)
Madrid, long degree						
Male	2,337	1,780	1,998	1,503	1,504	2,070
	(954)	(752)	(1,124)	(708)	(658)	(986)
Female	1,991	1,508	1,745	1,761	1,395	1,769
	(984)	(553)	(863)	(766)	(569)	(858)
Total	2,190	1,632	1,839	1,650	1,439	1,911
	(979)	(656)	(972)	(744)	(598)	(931)
Madrid, short degree						
Male	2,067	1,526	1,054		1,381	1,929
	(813)	(575)	(571)		(514)	(830)
Female	1,945	1,643	1,078		1,178	1,598
	(740)	(457)	(638)		(543)	(695)
Total	2,039	1,616	1,067		1,229	1,798
	(795)	(482)	(597)		(527)	(795)

Source: Own calculations from 'National Survey of Wage Structure', 2002.
Note: Standard deviations in parentheses.

and men. In the age range 25–29 years, females exhibit lower participation rates than men. The second source of selection bias arises from the fact that the Survey of Wage Structure reports wage earnings for employees, and therefore is restricted to those who decide to be wage earners and find a job. However, it is not possible to control for these sources of sample selection, as both participation decisions take place after graduation and may thus be conditional on events that take place after the survey. In any case, we use the data in Table 4 as a benchmark to evaluate expected wages of college students in our sample.

Analysis of the data in Table 4 reveals three findings about average earnings. First, they are greater for men than for women, with an aggregate gap at around 15 per cent. Second, average wages are generally higher the higher the degree duration. Third, employees in the Madrid region with degrees in S&E, social sciences, and humanities enjoy earnings that lie substantially above the national average. Differences in the cost of living and in industrial composition account for these differentials.

There is a positive gender wage gap between men and women for most degree types, with the exceptions of long humanities degrees (both in Spain and in Madrid) and short health degrees

Table 5. Sample distribution of college students in Madrid by degree type, year and duration of degree, and gender

	S&E	Health	Soc. Sci.	Hum.	Educ.	All
First year, long degree						
Male	108	5	40	34	7	194
Female	66	14	69	29	18	196
All	174	19	109	63	25	390
First year, short degree						
Male	22	4	10		2	38
Female	18	24	14		12	68
All	40	28	24		14	106
Junior, long degree						
Male	28	8	26	32	5	99
Female	19	18	42	38	14	131
All	47	26	68	70	19	230
Junior, short degree						
Male	29	3	7		3	42
Female	18	30	14		8	70
All	47	33	21		11	112

Source: Own calculations from 'Young people facing college education', 2001, 2004, and 2005.

in Madrid, for which the gender wage gap is negative. These degrees are associated with occupations that are dominated by women. Among the potential reasons for the gender gap, we should mention three: pure gender discrimination; the possibility that, with all other things equal, firm-specific accumulated human capital tends to be lower for women because they are more likely to experience discontinuities in their professional career; and occupational segregation. In the latter case, women are more likely to face restrictions that force them to choose occupations with lower wages in exchange for non-wage compensations, such as greater time flexibility. Regarding this, in those occupations for which the wage gap is reversed (related with long educational and short health degrees), women predominate.

We concentrate our analysis on two different subsamples of students, which correspond to extreme cases of the time to graduation, first-year and penultimate-year or junior students. We expect both their characteristics and wage expectations to differ very much for these two particular groups. In particular, we expect students closer to completion to have much lower uncertainty about their academic prospects, as well as a better informed assessment of their job market prospects after graduation.

In Table 5, we present the sample distribution of college students in our sample, for these two particular groups. In line with our earlier comments, our sample exhibits a low proportion of men in educational and health degrees of any duration, whereas S&E degrees, especially of long duration, are dominated by men.

3. Empirical analysis

3.1 Basic model

We use a stylized model of human capital accumulation and investment in education that suits the needs of our empirical analysis as in Webbink and Hartog (2004). For any individual,

we assume that his/her individual wage, W^*, is proportional to her amount of human capital, H. Assuming, without loss of generality, that unobserved individual factors are on average equal to 0, the expected log wage for a level of education S and a given set of observed individual factors equal to $\boldsymbol{Z}$, is $\alpha + \Phi S + \theta_1' Z$.

Moreover, for a university student in the k-th academic year of her college degree, her expected wage after graduation will depend on the information set determining her expectation. In particular,

$$E_k(lnW) = \alpha_k^e + \Phi_k^e S + \theta_{1k}^{e\prime} Z + E_k(v),$$

where $E_k(\bullet)$ represents the mathematical expectation, conditional on his/her information set, and α_k^e, Φ_k^e, θ_{1k}^e represent the expected returns in the wage equation of the corresponding variables in that information set. Assuming that $E_k(v)$ is equal to 0, then the expected average wage becomes $\alpha_k^e + \Phi_k^e S + \theta_{1k}^{e\prime} Z$.

Student information sets are related to the amount and quality of a student's knowledge about the economic value of his/her college degree, and to the time until receiving a wage as a graduate, i.e. his/her prediction horizon. Fresher students face much more uncertainty about their career prospects than junior students. First, their graduation probability is much lower for the first-year students, so that the effort that they put on computing an accurate forecast of their future wages is much lower. Second, the time horizon until they get into the labor market is much longer for first-year students. Third, first-year and junior students differ in their attitudes and their maturity. The first ones have finished secondary school very recently, whereas the last ones are very prone to graduating and getting into the labor market. Because gathering information is costly, first-year students are less willing to pay effort on learning about wages after graduation than junior students, who expect to be searching for a job in a much shorter time.[4] Fourth, the relative importance of the different aspects of the personal environment might differ substantially among first-year and junior students. In both cases, their personal environment is determined by their household and their college peers. In the case of first-year students, information is dominated by the one provided by the household, so that when forecasting wages they rely much more on the earnings opportunities of their parents. But because parents are in a very different stage of their life cycle, the information that provides about expected earnings may be less informative than the information that graduates in their first job can provide. Junior students are also influenced by their household. But the characteristics and attitudes of their peers are very different: they are more mature and much more concerned with their labor market prospects. Also, they are much closer to fresh graduates already working or searching for a job.

3.2 Empirical specification

In addition to the variables that characterize socioeconomic background and may be associated with human capital accumulated before higher education, it is also important to account for further individual characteristics. In particular, gender and the academic curriculum during secondary education may have a systematic effect on the subjective valuation of wages. Thus, we extend the vector of covariates, denoting it as $\boldsymbol{X}_i$. In addition to the unobservables affecting human capital obtained before higher education, there are individual characteristics that are unobserved in the data and affect subjective valuations. Also, because we will consider estimation for students in a particular course, the term will be part of the constant term. Therefore, we can write our empirical model as:

$$\ln W_i = \beta' X_i + u_i.$$

As we have mentioned in Section 2, we do not fully observe W_i, but a discretized version of it, W_i, which can be defined as:

$$W_i = j, if\ \mu_{j-1} < W_i < \mu_j\ (j = 1, \ldots, 5),$$

where the parameter values μ_j are known. Our estimation approach follows the strategy developed for models with multiple ordered responses that has been applied when using contingent-type data as, e.g., in Cameron and Quiggin (1994), Cai *et al.* (1998), and Papke (1998). In particular, our model is a pointwise censored model, as the thresholds determining the different categories are known, so there is no need to estimate them as parameters.[5] The model can be estimated by maximum likelihood after assuming a distribution for the unobservables. Besides, our knowledge of the thresholds allows us to obtain projections for expected wages as in a standard linear model.

4. Results

To assess the quality of wage forecasts by college students in Madrid, we exploit the information about expected wages reported by the students in our sample. Our estimates are subsequently used to compute individual predictions of expected wages and compare them with average actual wages for working graduates. The values reported for expected wages represent subjective valuations. This means that the estimated effects of the conditioning variables will combine the influence of these variables on the potential wage on the one hand, and the quality of the information used in computing wage expectations on the other. We have implemented separate maximum likelihood estimates for first-year and junior students. The corresponding results are shown in Table 6.

We consider two different specifications. The first one corresponds to the more general model, including all the selected variables, and the second one is the model of our choice, being the one that provides the best fit to the data. To achieve such specification, we test for the significance of several variables, both at the individual and at the joint levels, removing those that were clearly insignificant. We report the *p*-values of the corresponding tests. The goodness of fit of the estimated specifications is reasonably good. We will concentrate our comments on the results regarding our preferred specifications.

4.1 Determinants of expected wages

We first focus on the effects in wage expectations of two factors: gender and academic performance. We will discuss later the role of further variables in expectation formation.

4.1.1 The role of gender. The gender effect is significantly negative for first-year students, and its magnitude does not depend on the type of the degree, as the interaction between gender and degree is neither individually nor jointly significant. Hence, women realistically expect lower wages than men with similar characteristics. The fact that women present higher college attendance and better academic performance, as well as a greater expected probability of college graduation, is consistent with this result. Such characteristics propitiate a greater effort in information gathering, which results in lower overprediction of future wages. Nevertheless,

Table 6. Determinants of expected wages for college students in Madrid

	First-year students				Junior students			
	(i)		(ii)		(iii)		(iv)	
Gender	–0.207†	(0.091)	–0.117†	(0.059)	–0.004	(0.115)	–0.001	(0.114)
Repeater	0.033	(0.087)	0.062	(0.058)	0.249†	(0.118)	0.251†	(0.115)
Health × female	0.155	(0.221)			0.316	(0.252)	0.325	(0.259)
Educ. × female	0.012	(0.134)			–0.276	(0.169)	–0.310†	(0.155)
Soc.Sci. × female	0.207	(0.154)			–0.240	(0.170)	–0.241	(0.175)
Hum. × female	0.107	(0.132)			–0.169	(0.145)	–0.162	(0.144)
Health × repeater	–0.196	(0.202)			0.082	(0.251)	0.111	(0.238)
Educ. × repeater	–0.237	(0.215)			0.070	(0.201)	0.069	(0.198)
Soc. Sci. × repeater	0.246	(0.171)			–0.325	(0.216)	–0.321	(0.212)
Hum. × repeater	0.012	(0.133)			–0.508†	(0.160)	–0.495†	(0.150)
Health	–0.254	(0.209)	–0.206†	(0.090)	–0.402	(0.252)	–0.436*	(0.255)
Educ.	–0.398†	(0.127)	–0.475†	(0.120)	–0.063	(0.114)	–0.053	(0.109)
Soc.Sci.	–0.288†	(0.134)	–0.183*	(0.100)	0.052	(0.155)	0.048	(0.145)
Hum.	–0.285†	(0.119)	–0.290†	(0.099)	0.096	(0.121)	0.077	(0.109)
Public secondary	–0.019	(0.054)			0.001	(0.064)		
Access grade	0.001	(0.003)			0.003	(0.004)		
Access at first attempt	0.039	(0.075)			–0.031	(0.068)		
University father	0.113†	(0.056)	0.130†	(0.055)	0.053	(0.064)		
Science secondary	–0.205†	(0.079)	–0.223†	(0.079)	0.006	(0.084)		
Grant	–0.018	(0.068)			–0.063	(0.084)		
First choice	–0.058	(0.056)	–0.052	(0.052)	–0.144*	(0.074)	–0.142†	(0.072)
Same degree	–0.146*	(0.077)	–0.135*	(0.074)	0.010	(0.087)		
Economic indep.	0.047	(0.063)			0.034	(0.080)		
Parental influence	0.039	(0.060)			–0.098	(0.071)	–0.087	(0.068)
Vocation	–0.030	(0.069)			–0.001	(0.100)		
Difficulty	–0.110*	(0.061)	–0.096	(0.059)	0.030	(0.074)		
Satisfied	0.124*	(0.066)	0.101	(0.065)	0.006	(0.092)		
High family income	0.101	(0.067)	0.097	(0.065)	–0.060	(0.076)		
Short degree	–0.079	(0.069)	–0.091	(0.067)	–0.134*	(0.077)	–0.134*	(0.072)
Wald tests of group variables (% *p*-values)								
Reasons of choice	*34.6*				72.5			
Access grade	79.4				74.3			
Degree types	2.0		0.1		25.1		21.5	
Female × degree types	65.8				15.8		7.3	
Repeater × degree types	27.9				0.4		0.8	

Notes: Pointwise censored model without selection. We have controlled for the survey year included the corresponding binary variables. Standard errors in parentheses.
*,† Denote significance at 10 and 5 per cent, respectively.

the estimated magnitude does not offset the positive gender gap that is observed between men and women. This result suggests that women do not fully account for the gender gap in their future earnings.

Remarkably, unlike first-year students, the gender effect coefficient is not significant, but the differential gender effects across degrees are jointly significant. The differential gender effect for educational degrees is behind this joint significance. Given that we have considered S&E as reference group, this result implies that there are no differences in wage expectations by gender students in S&E degrees. For the remaining degree types, only women in health

degrees seem to expect higher wages than men, but the difference is not statistically significant. In fact, as shown in Table 4, there is a substantial negative gender gap against women in actual wages for S&E and health graduates. For the remaining degree types, there is a negative gap in wage expectations against women, but it is only significant in the case of educational degrees. Interestingly, the lowest gender gap in actual wages happens for education graduates, yet female education students appear the most pessimistic about their future wage prospects. It is worth noticing that in Spain the educational sector is highly feminized and mostly publicly owned. This might explain both the fact that salaries are similar between men and women, and the fact that these are significantly lower than the average.

4.1.2 The role of underperformance. We also control for underperformance at college through the variable repeater, for which we allow for interactions with the degree type. In the case of fresher students, being a repeater does not affect wage expectations after completion. We can thus conclude that underperforming at an early stage of college studies does not alter expectation formation.

Besides, we also find that fresher college students declaring to be satisfied with their studies tend to expect higher wages after completion. This variable is a subjective indicator of college satisfaction, which captures a different effect than measures of college performance, like repeater. Actually, the correlation between satisfied and repeater is below 6 per cent and clearly non-significant. The effect of this variable, though, is negligible in the case of experienced students.

The role of underperformance, as measured by the variable repeater, is only relevant for junior students and differs across degree types, as shown by the joint significance of the interaction terms. In line with Webbink and Hartog (2004), we find that, for most degrees, repeaters expect higher wages than non-repeaters. The joint significance of these interactions is mostly driven by the interaction term between humanities and repeater, which is clearly significant. This latter result is at odds with Webbink and Hartog (2004), but in line with Brunello *et al.* (2004), by which students who take longer than required to complete their degree have lower expected college wages.

We find, though, the opposite result for humanities students, and no differences between repeaters and non-repeaters in social sciences degrees. Our findings are compatible with Brunello *et al.* (2004), by which business and economics students who take longer than required to complete their degree have lower expected college wages.

4.1.3 Other factors. In addition to gender and performance, we also control for other factors. Most of them appear irrelevant for junior students.

Pre-university background, such as the grade achieved in the exam to access college, barely affects students' expected wages. As an exception, having followed a science curriculum in high school has a negative and significant effect on wage expectations for fresher students. Regarding the family background, living in a high-income family does not affect expectations. However, the fact that the father has a university degree tends to increase wage expectations for fresher students, but it is irrelevant for more experienced students.

In the case of junior students, neither pre-university nor household variables exhibit any effect on expected wages. These variables might have affected the choice of college degree undertaken, which can be related to the ability to predict expected wages shortly after the students have chosen and started their college degree. However, such variables do not affect expectation formation by junior students; to the extent that their degree completion is closer

and their (past) degree choice is distant. The only exception is the variable indicating whether the degree undertaken was the student's first choice, with significant negative effect. This result suggests that students who succeeded in entering their most preferred degree tend to expect lower wages.

With respect to the features behind the choice of college degree, we find different patterns among first-year and junior students. The fact that the student consistently applied to the same degree in different colleges exhibits a negative and significant effect only for first-year students. We interpret this result as that vocational students, who show strong preferences for a specific degree, tend to expect lower wages than those showing a higher taste for degree diversity. The fact that junior students are attending a degree which was their first choice leads to a negative effect on wage expectations, whereas it has no effect for first-year students.

We also find that the degree type, controlled by binary variables that are individually and jointly significant, does matter in expectation formation by first-year students. We observe that first-year students attending degrees different from S&E (the reference group) expect lower wages, especially in the case of humanities and, mostly, educational degrees. These results resemble the evidence reported in Table 4 about actual wages, by which college graduates in these latter degrees exhibit the lowest average actual wages. The results are substantially different, though, for junior students, for which, as mentioned earlier, differences across degrees for female and repeaters appear. Also, the degree duration (as captured by the variable *short* degree) positively affects wage expectations, but it is only significant for junior students.

4.2 Wage forecasts

We use our preferred estimates from Table 6 to predict student mean expected wages by gender and the degree type in Table 7. We also report, in Table 8, the percentage difference between mean wage expectations and the mean actual wages for young college graduates employees in Madrid (reported in Table 4). The comparison with the mean actual wages indicates that first-year students tend to overestimate their potential wages to a great extent.[6] The rankings of expected and actual wages by degree and gender, though, appear very similar. Namely, mean wages for long social sciences, and very specially S&E degrees, are generally among the highest, whereas the lowest mean wages usually correspond to educational degrees.

The level of overestimation is generally lower in the case of junior students, suggesting that students become more realistic as they approach graduation. This pattern has been pointed out by Betts (1996) and Smith and Powell (1990). They reach the conclusion that, in their case, college seniors may have more accurate information and form more realistic expectations. Brunello *et al*. (2004) claim that the negative correlation between college seniority and the expected wage identifies a learning effect, as the students become more realistic about their future earnings as they complete their curricula. They also consider alternative explanations to this effect: first, the possibility that senior students take the questionnaire more seriously, and second the fact that students might be taking wage inflation into account. We have no way to control for the first possibility, and our interviews were conducted in a period of price stability (inflation rate never grows above 4 per cent) and steady growth in the Spanish economy.

For each degree duration, the mean expected wages for junior students are very much alike among degrees, with the only exception of junior students in short educational degrees, for which mean expected wages are much lower.

With the exceptions of health and short educational degrees, expected wages for first-year female students are, on average, lower than those for men in the same group. However, looking at Table 8, considering that actual wages for graduates are systematically lower for

Table 7. Average expected wages for college students in Madrid by degree type, year and duration of degree, and gender

	S&E	Health	Soc.Sci.	Hum.	Educ.	Total
First course, long degree						
Male	4,433	2,878	3,393	2,790	1,861	3,798
	(807)	(634)	(830)	(574)	(520)	(1,082)
Female	3,821	3,113	3,324	2,430	1,458	3,173
	(629)	(612)	(938)	(722)	(607)	(1,037)
Total	4,201	3,051	3,350	2,625	1,571	3,484
	(800)	(609)	(896)	(666)	(602)	(1,103)
First course, short degree						
Male	3,416	2,839	2,752		1,556	3,083
	(966)	(464)	(643)		(39)	(936)
Female	3,114	2,853	2,776		2,040	2,763
	(826)	(786)	(791)		(595)	(833)
Total	3,280	2,851	2,766		1,971	2,878
	(907)	(742)	(718)		(575)	(881)
Penultimate course, long degree						
Male	2,971	2,278	2,709	2,567	2,301	2,682
	(476)	(390)	(405)	(448)	(517)	(488)
Female	3,124	2,720	2,049	1,919	1,713	2,224
	(461)	(569)	(385)	(555)	(534)	(673)
Total	3,033	2,584	2,302	2,215	1,868	2,421
	(471)	(553)	(506)	(601)	(580)	(641)
Penultimate course, short degree						
Male	2,392	1,834	2,469		1,940	2,333
	(525)	(265)	(205)		(453)	(493)
Female	2,452	2,289	1,923		1,185	2,131
	(515)	(371)	(182)		(471)	(550)
Total	2,415	2,247	2,105		1,391	2,207
	(516)	(383)	(322)		(566)	(536)

Source: Own calculations from 'Young people facing college education', 2001, 2004, and 2005.
Note: Standard deviations in parentheses.

women than for men, the relative level of overestimation is still higher for female than for male first-year students. Nevertheless, for each degree and duration, the gaps between expected and actual wages by gender are narrower when students are closer to graduation.

In fact, the differences in mean actual wages by gender and degree are higher than the corresponding differences in mean expected wages. Consequently, we observe substantial differences in the gap between mean expected wages and mean actual wages for junior students by degree and by gender. As extreme case, short degrees in social sciences exhibit the highest gap between mean expected and actual wages, particularly in the case of first-year students.

To understand these results, it must be noted that the actual wages by degree that we use as reference correspond to young college graduates who are employed in occupations that typically require the corresponding degree. Hence, when making the comparison between expected and actual wages, we are considering observed wages in specific occupations that are directly linked to the corresponding degree. In the case of educational degrees, we deduct from

Table 8. Percentage difference between average expected wages and average wages observed for working graduates in Madrid by degree type and duration, degree year, and gender

	S&E	Health	Soc.Sci.	Hum.	Educ.	Total
First year, long degree						
Male	89.7	61.6	69.9	85.6	23.7	83.5
Female	91.9	106.5	90.5	38.0	4.5	79.3
Total	91.8	87.0	82.1	59.1	9.2	82.3
First year, short degree						
Male	65.2	86.0	161.2		12.7	59.8
Female	60.1	73.7	157.4		73.2	72.9
Total	60.9	76.5	159.2		60.3	60.1
Junior, long degree						
Male	27.2	27.9	35.6	70.7	53.0	29.6
Female	56.9	80.4	17.4	9.0	22.8	25.7
Total	38.5	58.3	25.2	34.2	29.8	26.7
Junior, short degree						
Male	15.7	20.2	134.3		40.5	20.9
Female	26.1	39.4	78.3		0.6	33.4
Total	18.4	39.1	97.2		13.2	22.8

Source: Own calculations from 'Young people facing college education', 2001, 2004, and 2005, and 'National Survey of Wage Structure', 2002.

Table 9. Percentage difference in average expected wages for working graduates in Madrid between repeaters and non-repeaters by degree type and duration

	S&E	Health	Soc.Sci.	Hum.	Educ.	All
First-year, long degree	6.9	*–10.5*	–0.9	*–0.3*	*–2.2*	6.1
First-year, Short degree	30.8	*25.7*	–11.5		2.1	22.6
Junior, long degree	29.0	*27.1*	9.2	–20.1	*64.0*	20.5
Junior, short degree	9.1	35.4	*–6.4*		*67.6*	21.9

Source: Own calculations from 'Young people facing college education', 2001, 2004, and 2005.
Note: Those ratios with less than five observations in either group are in *italics*.

Table 4 that the occupations that are directly related to such degrees are worse paid than occupations associated to other college degrees. The higher gap between expected and actual wages can be partly explained by the fact that some students will consider aiming at other occupations that are not specific of educational degrees, which can provide them better salaries.

For first-year and junior students, the percentage difference between mean expected wages by degree type and duration, in accordance with being or not a repeater, is presented in Table 9. It must be noted that some cells have been computed with very few observations. We find that being a repeater is a source of distortion, to the extent that junior repeaters, in comparison with their remaining peers, still heavily overpredict. This result is in accordance with those in Webbink and Hartog (2004), by which repeaters wrongly tend to be more optimistic than non-repeaters.

4.3 Accuracy of expectations

A conclusion from our results is that as the time horizon toward graduation shortens, students' ability to realistically compute their expected wages improves. Junior students are less affected by pre-university or family background variables when computing their expected wages. Mostly, the type and the duration of the degree are the major variables affecting wage expectations of junior students. But also gender and college performance play a relevant role in their expectation formation. In the case of gender, female students realistically predict lower wages than their male counterparts. In the case of performance at college, repeaters tend to produce wage predictions of lower quality.

It must be noted that the overestimation of expected wages with respect to actual wages for young working graduates is actually greater than reflected in Table 8. This is because the individuals in our sample are not strictly comparable with the sample for which average actual wages were computed. This latter sample is restricted to young graduates who have decided to work and have indeed found a job as graduates. In contrast, our sample comprises students who have not yet graduated. For those who graduate, some will eventually not work, either because they decide not to enter the labor market or because they will not find a job. Moreover, a proportion of them will drop out of college before graduation. Therefore, it is possible that part of the apparent improvement in the formulation of expectations with increasing degree years merely reflects sample selection of students who are much more likely to work in jobs that require a university education. In this sense, our results are consistent with Betts (1996), who finds that students in higher years proved to be much better informed with respect to the labor market than fresher students.

We find that the level of accuracy in wage predictions differ substantially among students enrolled in different degrees. This phenomenon was also presented in Jerrim (2011). He finds that students on almost all disciplines significantly overestimate their starting salaries as college graduates, except for education students, who overestimate by a lesser amount, and medicine students, who tended to underestimate their expected wages with respect to actual wages. We do find the same pattern for educational degrees.

5. Conclusions

This paper models the wage forecasting of college students in Madrid universities. We use a microeconomic data set previously exploited by Alonso-Borrego *et al.* (2007) that includes academic, personal and household characteristics, as well as reported expected wages. This data set includes students from all universities and most degrees available in Madrid. This rich data set avoids concentrating on students in a particular type of college degree. Differences in time to completion may affect students' subjective valuation of college degrees. Such differences may affect individual processing of relevant information. For this reason, we consider two different subsamples, first-year and junior students.

The role of gender in wage expectations is particularly relevant. Women in their first year expect, other things equal, lower salaries than their male counterparts, and this gender gap is independent of the type of degree. Such gender differences persist for more experienced students, but differ by degree type. This evidence is consistent with the fact that women systematically earn lower mean wages than men. Also, college performance appears relevant in wage predictions for junior students: repeaters tend to overpredict wages to a larger extent than non-repeaters. We interpret this result as those students with worse college curriculum also yield wage predictions of lower quality.

We find that the degree type has also a relevant role on expected wages. There were also differences on how students produce their wage expectations depending on their degree year. Comparison of mean predicted expected wages with mean actual wages for young working graduates reveals a positive gap, which reflects that college students tend to overpredict their wages after graduation. This gap tends to narrow for junior students. This reflects that expectations became more realistic as students approach graduation. The information set is strongly influenced by the student's personal environment at the beginning of college studies. Such influence dilutes as students approach graduation and their uncertainty about their career prospects is reduced.

We find that students with gross underperformance have a distinct pattern when forecasting their future earnings in comparison with their peers. The existence of a clearly differentiated pattern might signal either a different expected carrier path or the presence of different information processing abilities.

Shortly after high school graduation, college students systematically overpredict their future wages. Even though first-year female students realistically predict lower wages than comparable men, their level of overprediction with respect to actual wages is higher. As college students approach graduation, their wage expectations, with some exception, become more realistic, and tend to be, on average, closer to the corresponding actual wages. It is important to remark that the breakdown by degree types shows differential patterns in accordance with the specific career prospects of each degree. We also find that students with a better defined career path exhibit wage expectations that keep closer to the observed wages, although still above.

We interpret overprediction, in line with Carvajal *et al*. (2000), as that student's expectations do not conform to market realities, to the extent that they usually expect to hold better job positions than what they actually will hold. This is because they have not properly understood the monetary value of their college education and they have chosen a path that implies either overqualification or training compatible with their expectations, but inefficient according market prices. The consequences of this overprediction are an excess supply of college graduates and an unsatisfactory match between the supply of college graduates and the labor market demand. This effect is greater among students who exhibit poor academic performance.

Notes

[1] The survey design is based on a nationwide data set produced jointly by the Centro de Investigaciones Sociológicas (National Sociological Institute) and the Ministry of Education in 1990, known as 'Los jóvenes ante la Universidad' ('Young people facing college education').

[2] Grants are awarded mostly for economic reasons, yet certain requirements on academic performance must be accomplished.

[3] We have disregarded self-employment status, for which nevertheless there are not reliable data sources about earnings. The proportion of young college graduates who were self-employed in Madrid, in accordance with the data from the 2002 Spanish Labor Force Survey, was below 10 per cent.

[4] Betts (1996) poses the existence of countervailing forces which make uncertain when information acquisition occurs more intensively. On the one hand, the marginal value of information may be greater in the early degree years, before high sunk costs make it costly to the student to change his/her career path. On the other hand, as far as information about the labor market acquired by the student does mainly come out from informal exchanges with peers, faculty, and others, then more experienced students might show an informational advantage over freshmen.

[5] An important practical advantage of exploiting wage thresholds by means of the pointwise censored model is that we do not need further assumptions about the distribution of the right tail to compute individual expected wages. More precisely, in a standard ordered probit in which the information on threshold values is not exploited, we must introduce an additional assumption for the right tail of the wage distribution (for declared expected monthly wages above 5,409 euro). Using results from the standard ordered probit estimates, we have found that predicted individual expected wages are very sensitive to this additional assumption.

[6] The only exception is for male students in short educational degrees, for which the mean expected wage is in line with the mean actual wage.

References

Alonso-Borrego C., Mora R. and Romero-Medina A. (2007) 'Una valoración de la enseñanza superior a partir de los salarios esperados y los salarios sombra de los estudiantes universitarios', *Cuadernos Económicos de ICE* 74: 11–38.

Betts J. R. (1996) 'What Do Students Know about Wages? Evidence from a Survey of Undergraduates', *The Journal of Human Resources* 31(1): 27–56. doi: 10.2307/146042.

Brunello G., Lucifora C. and Winter-Ebmer R. (2004) 'The Wage Expectations of European Business and Economics Students', *The Journal of Human Resources* 39(4): 1116–1142. doi: 10.3368/jhr.XXXIX.4.1116.

Cai Y., Deilami I. and Train K. (1998) 'Customer Retention in a Competitive Power Market: Analysis of a "Double-Bounded Plus Follow-Ups" Questionnaire', *The Energy Journal* 19(2): 191–215. doi: 10.5547/ISSN0195-6574-EJ-Vol19-No2-12.

Cameron T. A. and Quiggin J. (1994) 'Estimation Using Contingent Valuation Data from a Dichotomous Choice with Follow-up Questionnaire', *Journal of Environmental Economics and Management* 27(3): 218–234. doi: 10.1006/jeem.1998.1026.

Carvajal M. J. *et al*. (2000) 'Inter-gender Differentials between College Students' Earnings Expectations and the Experience of Recent Graduates', *Economics of Education Review* 19(3): 229–243. doi: 10.1016/S0272-7757(99)00034-5.

Chevalier A. (2007) 'Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates', *Oxford Bulletin of Economics and Statistics* 69(6): 819–842. doi: 10.1111/j.1468-0084.2007.00483.x.

Das M. and van Soest A. (1999) 'A Panel Data Model for Subjective Information on Household Income Growth', *Journal of Economic Behavior& Organization* 40(4): 409–426. doi: 10.1016/S0167-2681(99)00062-1.

Delaney L., Harmon C. and Redmond C. (2010) 'Decomposing gender differences in college student earnings expectations', University College Dublin Geary Institute Discussion Paper Series, Dublin, Ireland.

Dominitz J. and Manski C. F. (1996) 'Eliciting Student Expectations of the Returns to Schooling', *The Journal of Human Resources* 31(1): 1–26. doi: 10.2307/146041.

Jerrim J. (2011) 'Do UK Higher Education Students Overestimate Their Starting Salary?' *Fiscal Studies* 32(4): 483–509. doi: 10.1111/j.1475-5890.2011.00148.x.

Papke L. E. (1998) 'How are Participants Directing Their Participant-Directed Individual Account Pension Plans?' *The American Economic Review* 88(2): 212–216.

Smith H. L. and Powell B. (1990) 'Great Expectations: Variations in Income Expectations among College Seniors', *Sociology of Education* 63(3): 194–207.

Webbink D. and Hartog J. (2004) 'Can Students Predict Starting Salaries? Yes!' *Economics of Education Review* 23(2): 103–113. doi: 10.1016/S0272-7757(03)00080-3.